INSIGHT

In Search of the Molecular Basis of Heterosis

Heterosis, or hybrid vigor, refers to the phenomenon that progeny of diverse inbred varieties exhibit greater biomass, speed of development, and fertility than the better of the two parents (Figure 1). This phenomenon has been exploited extensively in crop production and has been a powerful force in the evolution of plants. The genetic basis has been discussed for nearly a century (Shull, 1908; Bruce, 1910; Jones, 1917), but little consensus has emerged. With the advent of the genomic era, the tools to establish a molecular basis for heterosis are at hand. In the past, there has been a tendency to attribute any molecular differences between the parents and progeny as contributing to the basis of heterosis. Some individuals dismiss the phenomenon as hopelessly complex. It seems likely, however, that the complexity derives from its multigenic nature and that eventually a unifying principle will emerge. In this article, we summarize some of the salient features of heterosis that a viable molecular model must explain.

The classic quantitative genetic explanations for heterosis center on two concepts (Crow, 1948). The first is "dominance," which originally meant that heterosis results from the complementation in the hybrid of different deleterious alleles that were present in the inbred parental lines by superior alleles from the opposite parent. Over time, this term came to mean the degree to which the heterozygous genotype performs differently from the mean of the two homozygous classes. The second historical explanation for heterosis is "over-dominance," which refers to the idea that allelic interactions occur in the hybrid such that the heterozygous class performs better than either homozygous class. Although these terms have developed a following in each case, they both now refer to nonadditive situations, differing in degree. These terms were coined before the molecular concepts of genetics were formu-

Figure 1. Heterosis in Maize.

Representative individuals from two inbred maize lines (B73 [far left] and Mo17 [far right]) and the progeny of reciprocal hybrid crosses (B73/Mo17 [left center] and Mo17/B73 [right center]; the female parent is listed first in maize genetics nomenclature) are shown. B73 and Mo17 are two high-quality inbred lines. Nonetheless, the progeny of a hybrid cross between these two lines are taller and more productive than either parent, illustrating the concept of heterosis.

lated and are not connected with molecular principles. Therefore, they are of diminished utility for describing the molecular parameters that accompany heterosis.

At the molecular level, one can envision two extreme models to explain heterosis. In the first case, one might imagine that in the hybrid, when the two different alleles of various genes are brought together, there is a combined allelic expression. In the second model, the combination of differ-

ent alleles produces an interaction that causes gene expression in the hybrid to deviate relative to the midparent predictions (e.g., by an upregulation of many housekeeping genes). This scenario might best be considered the result of regulatory gene allelic interactions. Indeed, a recent article by Song and Messing (2003), described in more detail below, provides evidence for altered regulatory effects in hybrids. The challenge in the development of

a molecular model for heterosis is to make the correct associations between phenotype and any causative molecular events that occur in hybrids.

The preferred explanation for heterosis during the past century was that the two inbred lines contain different slightly deleterious alleles at multiple loci. When the hybrid is produced, all mutations are complemented, generating progeny that exceed either parent. An early criticism of this idea was that if this hypothesis were true, it should be possible to create an inbred line with all of the superior alleles, showing little or no hybrid vigor—a situation that has not occurred. The counter argument stated that, with so many genes involved, it would be impossible to accumulate all of the better alternatives into one line as a result of the linkage of the deleterious alleles with superior alleles of other genes. Although it is true that any deleterious alleles would become homozygous in different inbred lines and that the hybrid would show complementation for these genes, this fact might account only for the hybrid being equivalent to the better of the two parents for the effect of any individual gene. Alternatively, if the complementation of alleles in different genes were cumulative in the phenotype, then heterosis would result. The molecular question to be addressed is whether the simple complementation of different slightly deleterious alleles generates a growth response that can explain heterosis. However, several observations regarding heterosis suggest that the basic principle of heterosis is something other than simple complementation.

The first observation to suggest a basis other than simple complementation is that although inbred lines have been improved greatly over the decades, the magnitude of heterosis has not diminished but has increased slightly (East, 1936; Duvick, 1999). If heterosis were caused by the complementation of deleterious alleles and inbred lines have been purged of the most severe of such alleles, then the absolute amount of heterosis might be expected to decline somewhat. Selection for better inbred lines

has improved the vigor of plants, but it has done less to change the magnitude of heterosis. In other words, heterosis gives the appearance of being more resistant to artificial selection than the quality of inbred lines themselves. Furthermore, the quality of two inbred lines does not necessarily predict the amount of heterosis; this must be determined in a cross. These observations suggest that instead of replacing alleles of genes that modulate physiological processes important for heterosis, the slight increase in hybrid vigor over the years might have occurred by selecting alleles at the right set of loci that produce the best combinations in hybrids to bring about heterosis.

A second observation about heterosis that argues against simple complementation is progressive heterosis in tetraploids (Levings et al., 1967; Mok and Peloquin, 1975; Groose et al., 1989; Bingham et al., 1994). At the diploid level, only two alleles of a gene can occur in an individual, but at higher ploidy levels, a variety of allelic combinations are possible for any one gene. In autotetraploids that are hybrid between two inbred lines (alleles designated AABB), heterosis occurs, but it is typically greater when there are potentially three or four different alleles present at the various loci (designated ABCD). Even in allohexaploid wheat, in which three different genomes contribute to the genetic constitution, hybrids between diverse varieties exhibit heterosis (Briggle, 1963). It appears that vigor improves with the greater number of distinct genomes present. For simple complementation to explain progressive heterosis, each new step-wise combination of genomes would need to supply increasingly superior alleles to complement the preexisting rate-limiting alleles without introducing deleterious alleles at other loci. The probability that this situation would occur is quite low. A release from a negative dosage effect on vigor by identical alleles could account for progressive heterosis, as discussed further below.

The third observation to consider is that inbreeding depression (the decline in vigor

over several generations of selfing) in tetraploids of many species proceeds faster than expected based on the homozygosis of alleles (Randolph, 1942; Alexander and Sonnemaker, 1961; Busbice and Wilsie, 1966; Rice and Dudley, 1974). In a diploid, selfing of a heterozygote (A/B) will produce half of the progeny that are homozygous at any one locus and the other half that regenerate the heterozygous situation. In an autotetraploid, the selfing of a heterozygote (A/A/B/B) will produce homozygotes (A/A/A/A or B/B/B/B) at any one locus in only $\sim$1 of 18 offspring (depending on the degree of linkage to the centromere). In addition to A/A/B/B heterozygotes being formed again, A/A/A/B and B/B/B/A heterozygotes are present in the population. Despite this difference in the rate of progression to homozygosity, the trajectory of inbreeding depression in tetraploids often is faster than predicted and not very different from that in diploids. In some species, tetraploid inbreeding depression proceeds faster than at the diploid level. As discovered by Randolph (1942), tetraploid derivatives of inbred maize lines often are less vigorous than the diploid progenitor. Thus, in this species, the end product of inbreeding depression in tetraploids is less than that of diploids, even though the genotype is identical (but differs in dosage). One resolution of this finding is to suggest that allelic dosage plays a more important role in tetraploids for generating inbreeding depression than does complete homozygosis itself, because the allelic dosage shifts more rapidly than homozygosis during selfing. The increasing number of identical alleles appears to have a negative dosage effect on vigor.

If there is any involvement of a dosage effect of alleles in polyploid heterosis, this realization is gratifying given that the bulk of quantitative trait loci show some degree of semidominant behavior (Tanksley, 1993), indicating that the control of quantitative characters is largely affected by multiple loci that exhibit an allelic dosage effect. The results of aneuploid studies likewise indicate that quantitative characteristics are affected by multiple dosage-

INSIGHT

dependent genes (Lee et al., 1996). There is likely a connection between these two observations (Guo and Birchler, 1994).

What is responsible for such dosage effects? It has been argued that these dosage effects are a reflection of dosage-dependent gene regulatory hierarchies (Birchler et al., 2001). Most regulatory genes exhibit some measure of dosage dependence, whereas target housekeeping genes usually show a greater degree of dominance/recessive behavior between allelic alternatives (Birchler and Auger, 2003). One possible explanation for this partial dichotomy comes from an analysis of dosage-sensitive genes in yeast (Papp et al., 2003). Loci that tend to have a significant haplo-insufficient effect on growth in diploid yeast encode products that are involved in molecular complexes. Regulatory genes in multicellular organisms often function as part of complexes, so if the same rules apply, regulatory genes usually will exhibit some measure of dosage dependence, whereas genes that encode metabolic functions will be less likely to show a dosage effect. Empirical observations indicate that most regulatory genes do exhibit some type of dosage response (Birchler et al., 2001). Consequently, quantitative traits will be controlled in large part by multiple dosage-dependent regulatory loci. Following this train of thought, one is led to the idea that heterosis is the result of different alleles being present at loci that contribute to the regulatory hierarchies that control quantitative traits.

Are there any data available on gene expression in inbreds and hybrids to suggest a shift in gene regulation in hybrids? A recent study by Song and Messing (2003), as well as work in our own laboratory (Osborn et al., 2003) and earlier work at the protein level by Romagnoli et al. (1990) and Leonardi et al. (1991), indicate that the expression of many genes does not exhibit the expected midparent value. In the study by Song and Messing, zein gene expression was studied in the endosperm of two inbred lines and their reciprocal hybrids. There are many zein genes that contribute to the total storage protein pool. Zein

cDNAs were cloned by reverse transcriptase–mediated PCR and then sequenced in large numbers. The relative expression of the various zein genes then were determined by the relative frequency of cDNAs for the various genes present in the sequenced sample. Of the 10 genes studied, only in one case did the hybrid expression follow the predictions of allelic dosage contributing to the genotype. Remarkably, in this study and those cited above, the range of deviation fell within either a twofold increase or a twofold decrease. The question then becomes, are these changes responsible for heterosis or a result of it? And if they are responsible, what property of heterozygosity at regulatory loci would produce this response at the target genes?

It is interesting to compare these results with the behavior of enhancer trap lines in species hybrids in the genus *Drosophila*. The changes in zein gene expression in maize hybrids have parallels with the altered reporter gene activity found in species hybrids. Hammerle and Ferrus (2003) examined the expression of enhancer trap lines in *D. melanogaster* and in hybrids with *D. simulans* or *D. mauritiana* using *lacZ* as a reporter. In three different homozygous strains of *D. melanogaster*, no variation in expression was found. However, in the interspecific hybrids, *lacZ* expression was modulated either up or down relative to that in the *D. melanogaster* control. Thus, the behavior of target genes in this case shows a pattern similar to that found by Song and Messing (2003). These findings indicate that bringing together divergent regulatory hierarchies will cause global modulations of target gene expression.

The magnitude of the changes in plant hybrids is interesting given that heterotic characteristics seldom show greater than twofold effects in terms of the change in biomass or fertility, although spectacular exceptions exist. Nevertheless, it is important to remember that even trivial changes in fertility compounded at each generation will produce a tremendous advantage for the favorable genotype. Those genotypes that maintain heterozygosity, and hence

heterosis, will have an advantage over alternative genotypes under most conditions. Thus, documenting these small changes in gene expression will help us understand large effects over evolutionary time.

If changes in gene expression are responsible for heterosis, which genes are involved? Furthermore, how do these changes compare with the alterations in gene expression that occur in aneuploids, which in most cases are detrimental to vigor? Aneuploidy also causes changes in gene expression typically within a twofold range (Birchler, 1979; Birchler and Newton, 1981; Guo and Birchler, 1994; Auger et al., 2001; Wanous et al., 2003). These changes can result from structural gene dosage effects, but more often they result from *trans*-acting effects that modulate the expression of most of the genome (Birchler et al., 2001; Matzke et al., 2003). It has been proposed that the reductions in gene expression that occur in both monosomics and trisomics are rate limiting on the phenotype and therefore act as underlying contributors to aneuploid syndromes (Birchler and Newton, 1981; Guo and Birchler, 1994; Birchler et al., 2001). It appears that the reductions in gene expression are detrimental to the vigor of the aneuploid plants. To date, these analyses have relied on a sampling of gene expression rather than a comprehensive examination of genome-wide expression patterns. A larger sampling might determine, for example, if heterosis in general is correlated with a majority of the increases in gene expression while aneuploidy leads to a significant number of reductions in gene expression in both monosomics and trisomics. A more complete picture might elucidate this distinction, if there is a meaningful comparison to be made. What distinguishes the phenotypic consequences of the ups and downs of gene expression in aneuploids versus hybrids? One possibility is that the gene expression changes that foster increased biomass and fertility have been selected in hybrid states over long periods of time, whereas aneuploid situations usually are transitory and the result of laboratory manipulations.

The Plant Cell

Microarray analyses might be able to provide some answers about the spectrum of genes that change in expression in hybrids and in which direction the changes occur. These studies, however, typically discard alterations in gene expression that are below a twofold cutoff. It is possible, even probable, that the truly relevant changes would be obscured by this type of data treatment. However, as the technology and statistical analysis of microarray data improve, the ability to go below this threshold will improve.

To formulate a molecular model of heterosis, simple broad alternatives need to be tested so that more refined and targeted hypothesis testing can focus on the detailed mechanism. One could argue that nothing less than defining how the genome interacts to create the phenotype is needed for an understanding of heterosis and that this understanding is too far in the future to attempt any examination of heterosis at present. Such a view is too agnostic and should not stand in the way of chipping away at alternatives. An eventual molecular explanation of heterosis will determine whether it can be manipulated for the benefit of agriculture and biotechnology.

James A. Birchler
Donald L. Auger
Nicole C. Riddle
Division of Biological Sciences
University of Missouri
Columbia, MO 65211
birchlerj@missouri.edu

REFERENCES

Alexander, D.E., and Sonnemaker, E.H. (1961). Inbreeding depression in autotetraploid maize. Maize Genet. Coop. Newsl. **35,** 45.

Auger, D.L., Newton, K.J., and Birchler, J.A. (2001). Nuclear gene dosage effects upon the expression of maize mitochondrial genes. Genetics **157,** 1711–1721.

Bingham, E.T., Groose, R.W., Woodfield, D.R., and Kidwell, K.K. (1994). Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. Crop Sci. **34,** 823–829.

Birchler, J.A. (1979). A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. Genetics **92,** 1211–1229.

Birchler, J.A., and Auger, D.L. (2003). Biological consequences of dosage dependent gene regulatory mechanisms in multicellular eukaryotes. In Biology of Dominance, R. Veitia, ed (Georgetown, TX: Landes Bioscience).

Birchler, J.A., Bhadra, U., Pal Bhadra, M., and Auger, D.L. (2001). Dosage dependent gene regulation in multicellular eukaryotes: Implications for dosage compensation, aneuploid syndromes and quantitative traits. Dev. Biol. **234,** 275–288.

Birchler, J.A., and Newton, K.J. (1981). Modulation of protein levels in chromosomal dosage series of maize: The biochemical basis of aneuploid syndromes. Genetics **99,** 247–266.

Briggle, L.W. (1963). Heterosis in wheat: A review. Crop Sci. **3,** 407–412.

Bruce, A.B. (1910). The Mendelian theory of heredity and the augmentation of vigor. Science **32,** 627–628.

Busbice, T.H., and Wilsie, C.P. (1966). Inbreeding depression and heterosis in autotetraploids with application to *Medicago sativa* L. Euphytica **15,** 52–67.

Crow, J.F. (1948). Alternative hypotheses of hybrid vigor. Genetics **33,** 477–487.

Duvick, D.N. (1999). Heterosis: Feeding people and protecting natural resources. In Genetics and Exploitation of Heterosis in Crops, J.G. Coors and S. Pandey, eds (Madison, WI: American Society of Agronomy/Crop Science Society of America), pp. 19–29.

East, E.M. (1936). Heterosis. Genetics **21,** 375–397.

Groose, R.W., Talbert, L.E., Kojis, W.P., and Bingham, E.T. (1989). Progressive heterosis in autotetraploid alfalfa: Studies using two types of inbreds. Crop Sci. **29,** 1173–1177.

Guo, M., and Birchler, J.A. (1994). Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science **266,** 1999–2002.

Hammerle, B., and Ferrus, A. (2003). Expression of enhancers is altered in *Drosophila melanogaster* hybrids. Evol. Dev. **5,** 221–230.

Jones, D.F. (1917). Dominance of linked factors as a means of accounting for heterosis. Genetics **2,** 466–479.

Lee, E.A., Darrah, L.L., and Coe, E.H. (1996). Dosage effects on morphological and quantitative traits in maize aneuploids. Genome **39,** 898–908.

Leonardi, A., Damerval, C., Herbert, Y., Gallais, A., and de Vienne, D. (1991). Association of protein amount polymorphisms (PAP) among maize lines with performances of their hybrids. Theor. Appl. Genet. **82,** 552–560.

Levings, C.S., Dudley, J.W., and Alexander, D.E. (1967). Inbreeding and crossing in autotetraploid maize. Crop Sci. **7,** 72–73.

Matzke, M.A., Mette, M.F., Kanno, T., and Matzke, A.J.M. (2003). Does the intrinsic instability of aneuploid genomes have a causal role in cancer? Trends Genet. **19,** 253–256.

Mok, D.W.S., and Peloquin, S.J. (1975). Breeding value of 2n pollen (diploandroids) in tetraploid × diploid crosses in potato. Theor. Appl. Genet. **46,** 307–314.

Osborn, T.C., Pires, J.C., Birchler, J.A., Auger, D.L., Chen, Z.J., Lee, H.-S., Comai, L., Madlung, A., Doerge, R.W., Colot, V., and Martienssen, R.A. (2003). Understanding mechanisms of novel gene expression in polyploids. Trends Genet. **19,** 141–147.

Papp, B., Pal, C., and Hurst, L.D. (2003). Dosage sensitivity and the evolution of gene families in yeast. Nature **424,** 194–197.

Randolph, L.F. (1942). The influence of heterozygosis on fertility and vigor in autotetraploid maize. Genetics **27,** 163.

Rice, J.S., and Dudley, J.W. (1974). Gene effects responsible for inbreeding depression in autotetraploid maize. Crop Sci. **14,** 390–393.

Romagnoli, S., Maddaloni, M., Livini, C., and Motto, M. (1990). Relationship between gene expression and hybrid vigor in primary root tips of young maize (*Zea mays* L.) plantlets. Theor. Appl. Genet. **80,** 769–775.

Shull, G.H. (1908). The composition of a field of maize. Am. Breeders Assoc. Rep. **4,** 296–301.

Song, R., and Messing, J. (2003). Gene expression of a gene family in maize based on non-collinear haplotypes. Proc. Natl. Acad. Sci. USA **100,** 9055–9060.

Tanksley, S.D. (1993). Mapping polygenes. Annu. Rev. Genet. **27,** 205–233.

Wanous, M.K., Munkvold, J.D., Kruse, J.D., Brachman, E.E., Klawiter, M.A., and Fuehrer, K.J. (2003). Identification of chromosome arms influencing expression of the HMW glutenins in wheat. Theor. Appl. Genet. **106,** 213–220.